FIG. 1

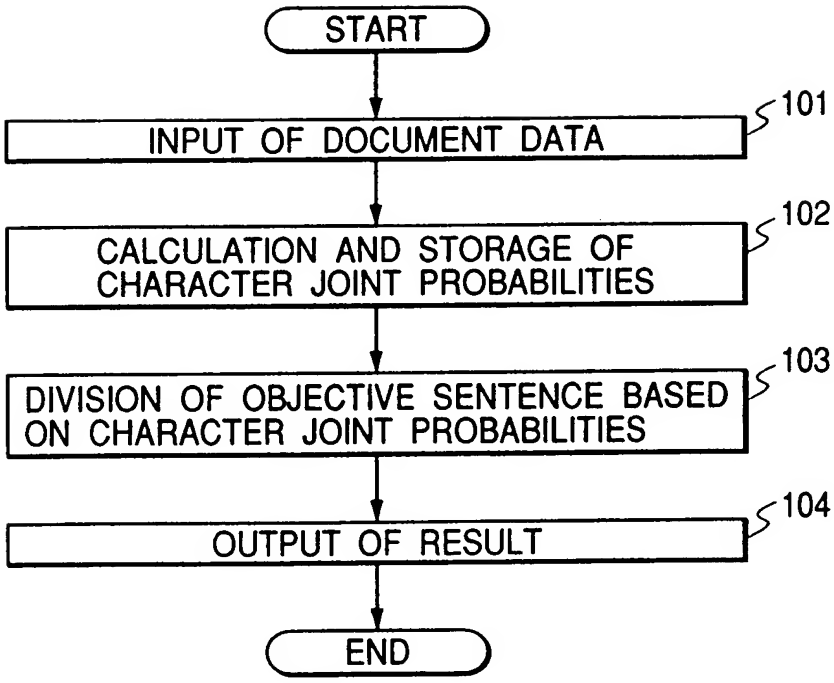

START
INPUT OF DOCUMENT DATA
101
CALCULATION AND STORAGE OF CHARACTER JOINT PROBABILITIES
102
DIVISION OF OBJECTIVE SENTENCE BASED ON CHARACTER JOINT PROBABILITIES
103
OUTPUT OF RESULT
104
END

FIG. 2

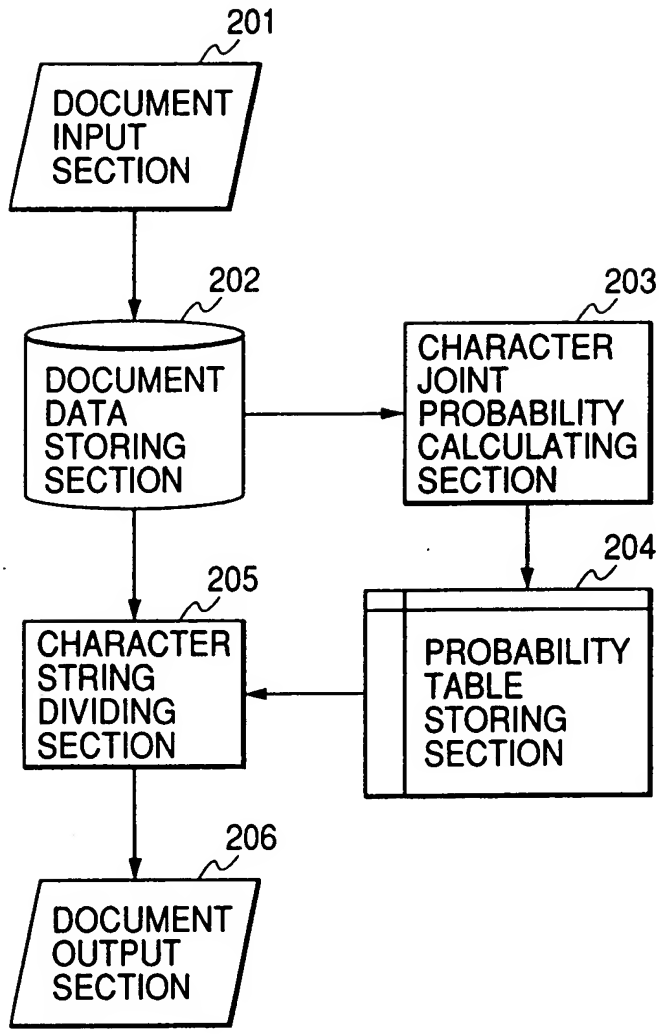

201
DOCUMENT INPUT SECTION
202
DOCUMENT DATA STORING SECTION
203
CHARACTER JOINT PROBABILITY CALCULATING SECTION
204
PROBABILITY TABLE STORING SECTION
205
CHARACTER STRING DIVIDING SECTION
206
DOCUMENT OUTPUT SECTION

FIG. 3

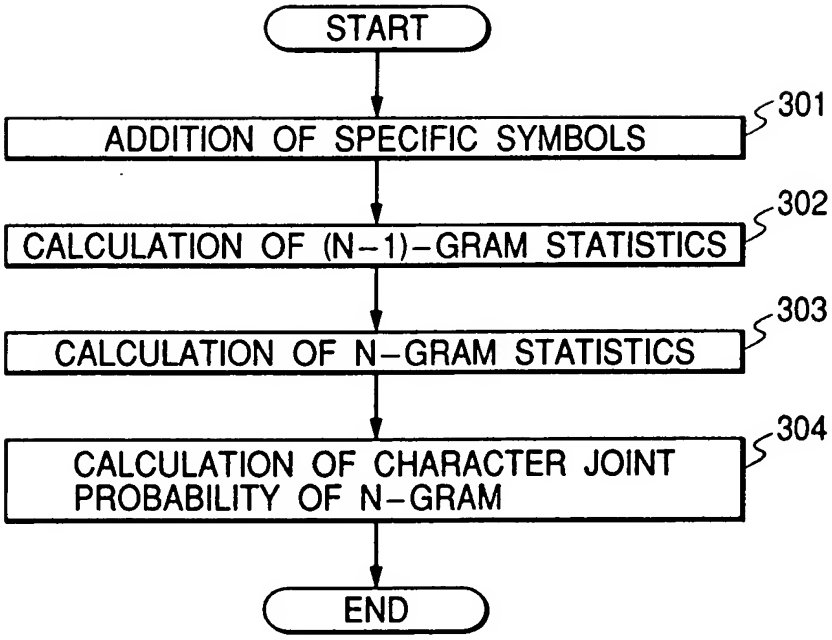

START
ADDITION OF SPECIFIC SYMBOLS
301
CALCULATION OF (N-1)-GRAM STATISTICS
302
CALCULATION OF N-GRAM STATISTICS
303
CALCULATION OF CHARACTER JOINT PROBABILITY OF N-GRAM
304
END

FIG. 4A

GIVEN CHARACTER STRING : a b a a b a

↓

OBJECTIVE CHARACTER STRING : # a b a a b a

FIG. 4B

APPEARANCE FREQUENCY OF 2-GRAM	
CHARACTER STRING	NUMBER OF TIMES
#a	1
ab	2
ba	2
aa	1

FIG. 4C

APPEARANCE FREQUENCY OF 3-GRAM	
CHARACTER STRING	NUMBER OF TIMES
#ab	1
aba	2
baa	1
aab	1

FIG. 4D

CHARACTER JOINT PROBABILITY	
CHARACTER STRING	PROBABILITY
#ab	1.0
aba	1.0
baa	0.5
aab	1.0

CALCULATION METHOD

=COUNT(#ab)/COUNT(#a)
=COUNT(aba)/COUNT(ab)
=COUNT(baa)/COUNT(ba)
=COUNT(aab)/COUNT(aa)

FIG. 4E

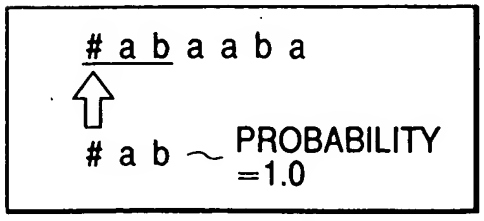


FIG. 4F

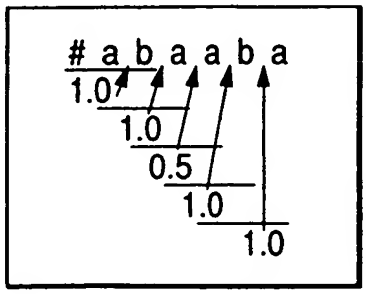


FIG. 5

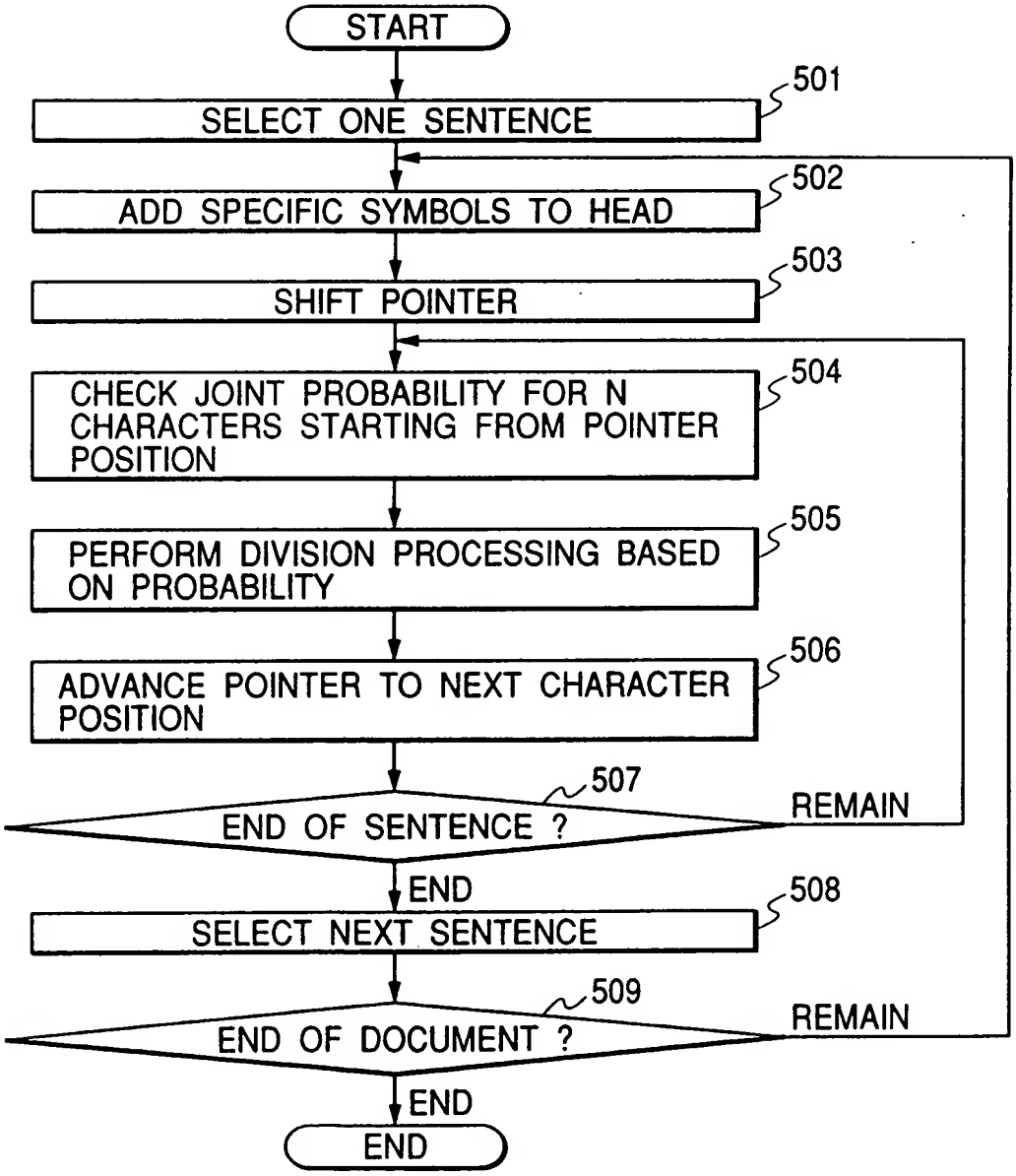

START
SELECT ONE SENTENCE
501
ADD SPECIFIC SYMBOLS TO HEAD
502
SHIFT POINTER
503
CHECK JOINT PROBABILITY FOR N CHARACTERS STARTING FROM POINTER POSITION
504
PERFORM DIVISION PROCESSING BASED ON PROBABILITY
505
ADVANCE POINTER TO NEXT CHARACTER POSITION
506
END OF SENTENCE ?
507
REMAIN
END
SELECT NEXT SENTENCE
508
END OF DOCUMENT ?
509
REMAIN
END
END

FIG. 6A

GIVEN CHARACTER STRING : にわにわにわ

↓

OBJECTIVE CHARACTER STRING : # にわにわにわ

FIG. 6B

APPEARANCE FREQUENCY OF 2-GRAM	
CHARACTER STRING	NUMBER OF TIMES
# に	1
にわ	3
わに	2

FIG. 6C

APPEARANCE FREQUENCY OF 3-GRAM	
CHARACTER STRING	NUMBER OF TIMES
# にわ	1
にわに	2
わにわ	2

FIG. 6D

CHARACTER JOINT PROBABILITY	
CHARACTER STRING	PROBABILITY
# にわ	1.0
にわに	0.6
わにわ	1.0

CALCULATION METHOD

=COUNT(# にわ)/COUNT(# に)
=COUNT(にわに)/COUNT(にわ)
=COUNT(わにわ)/COUNT(わに)

FIG. 6E

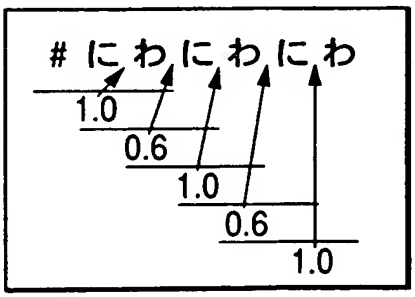


FIG. 7

CHARACTER JOINT PROBABILITY	
CHARACTER STRING	PROBABILITY
ようち	0.00016
ようっ	0.00011
ようつ	0.00005
ようて	0.00011
ようで	0.01938
ようと	0.08398
ような	0.13665
ように	0.22713
よくあ	0.01593
よくい	0.00612
よくえ	0.00024
よくお	0.00269
よくか	0.00171

FIG. 8

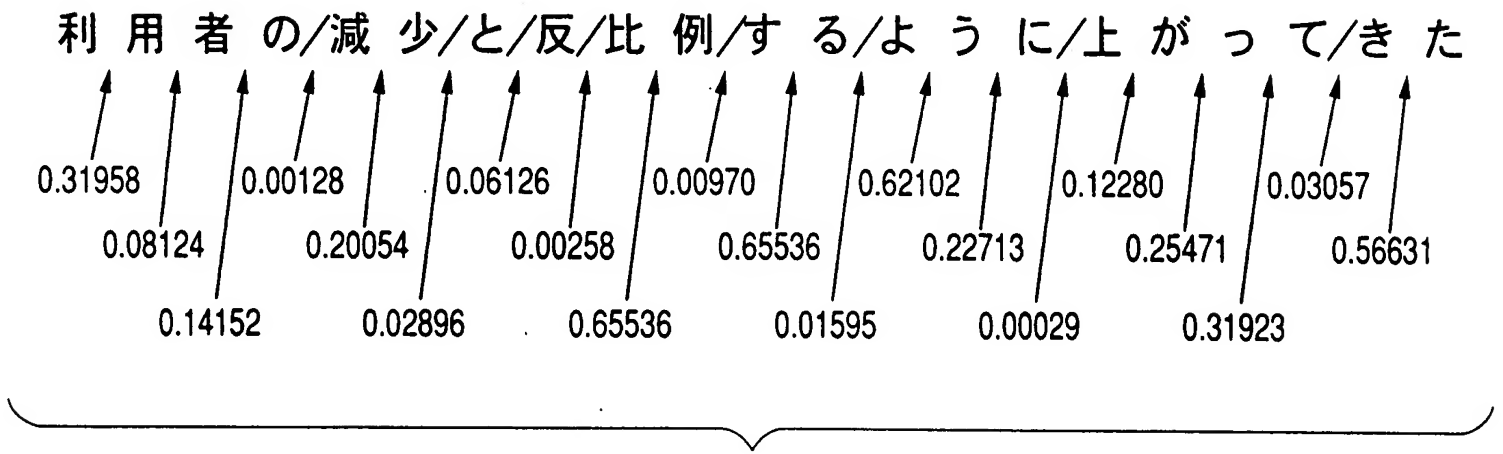

利 用 者 の/減 少/と/反/比 例/す る/よ う に/上 が っ て/き た
0.31958
0.08124
0.14152
0.00128
0.20054
0.02896
0.06126
0.00258
0.65536
0.00970
0.65536
0.01595
0.62102
0.22713
0.00029
0.12280
0.25471
0.31923
0.03057
0.56631

FIG. 9

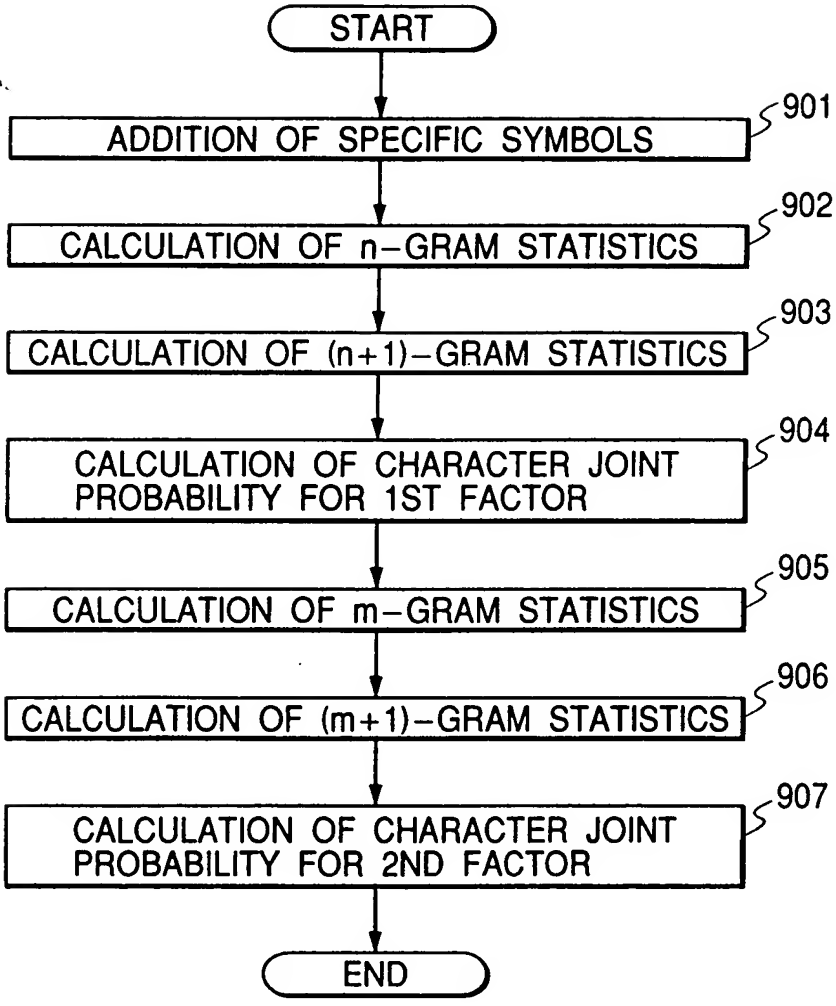

START
ADDITION OF SPECIFIC SYMBOLS
901
CALCULATION OF n-GRAM STATISTICS
902
CALCULATION OF (n+1)-GRAM STATISTICS
903
CALCULATION OF CHARACTER JOINT PROBABILITY FOR 1ST FACTOR
904
CALCULATION OF m-GRAM STATISTICS
905
CALCULATION OF (m+1)-GRAM STATISTICS
906
CALCULATION OF CHARACTER JOINT PROBABILITY FOR 2ND FACTOR
907
END

FIG. 10

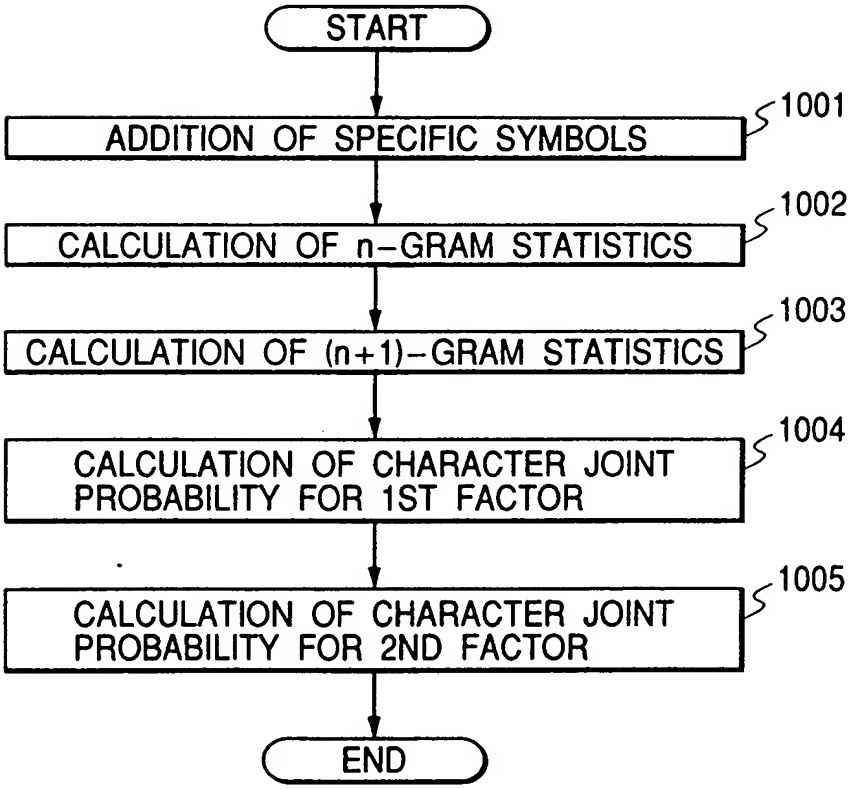

START
ADDITION OF SPECIFIC SYMBOLS
1001
CALCULATION OF n-GRAM STATISTICS
1002
CALCULATION OF (n+1)-GRAM STATISTICS
1003
CALCULATION OF CHARACTER JOINT PROBABILITY FOR 1ST FACTOR
1004
CALCULATION OF CHARACTER JOINT PROBABILITY FOR 2ND FACTOR
1005
END

FIG. 11

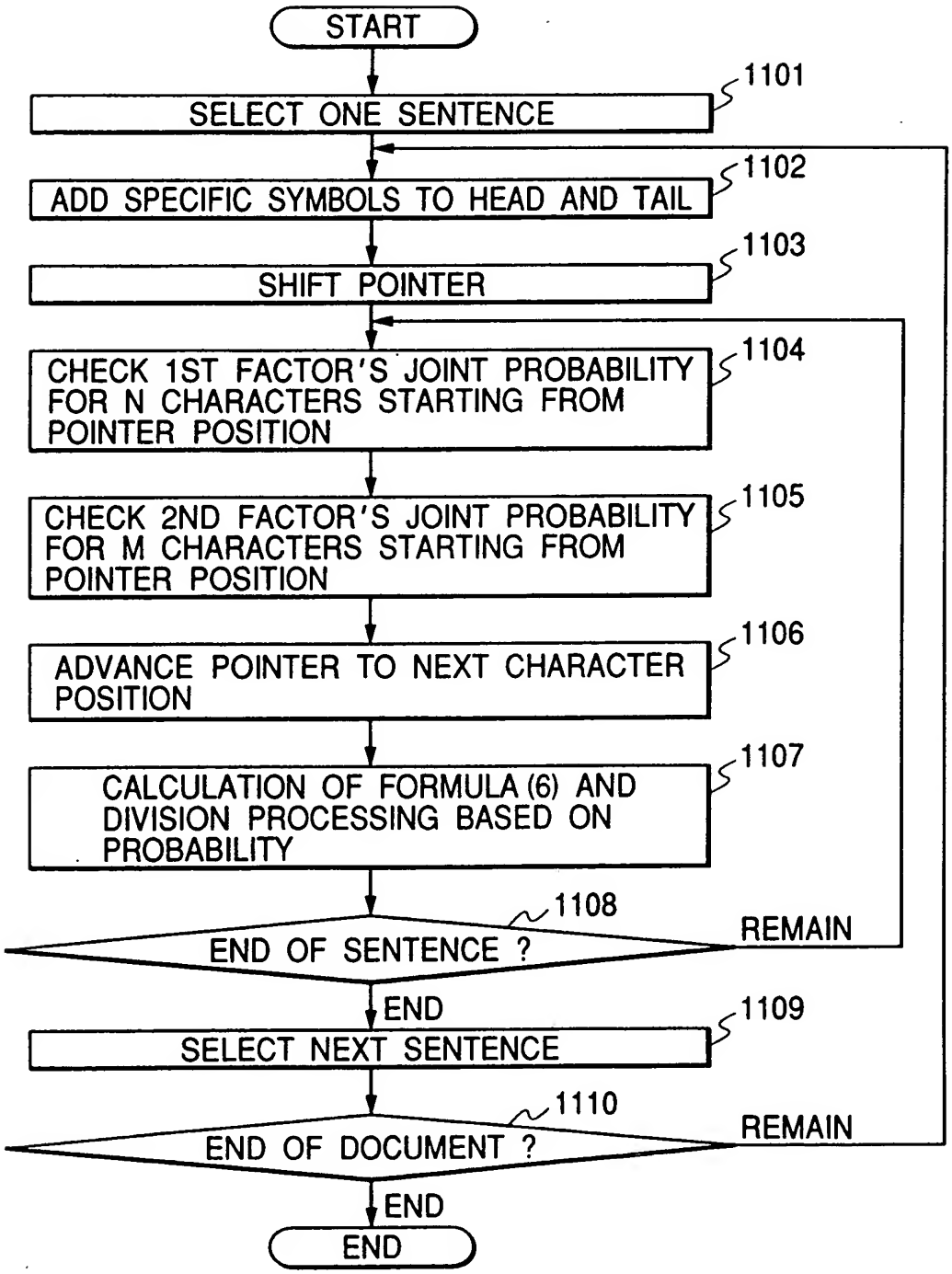

START
SELECT ONE SENTENCE
1101
ADD SPECIFIC SYMBOLS TO HEAD AND TAIL
1102
SHIFT POINTER
1103
CHECK 1ST FACTOR'S JOINT PROBABILITY FOR N CHARACTERS STARTING FROM POINTER POSITION
1104
CHECK 2ND FACTOR'S JOINT PROBABILITY FOR M CHARACTERS STARTING FROM POINTER POSITION
1105
ADVANCE POINTER TO NEXT CHARACTER POSITION
1106
CALCULATION OF FORMULA (6) AND DIVISION PROCESSING BASED ON PROBABILITY
1107
END OF SENTENCE ?
1108
REMAIN
END
SELECT NEXT SENTENCE
1109
END OF DOCUMENT ?
1110
REMAIN
END
END

FIG. 12A

GIVEN CHARACTER STRING : 仕事は仕事

↓

OBJECTIVE CHARACTER STRING : # 仕事は仕事 #

FIG. 12B

APPEARANCE FREQUENCY OF 2-GRAM	
CHARACTER STRING	NUMBER OF TIMES
# 仕	1
仕事	2
事は	1
は仕	1
事 #	1

FIG. 12D

APPEARANCE JOINT PROBABILITY		
CHARACTER STRING	1ST PROBABILITY	2ND PROBABILITY
# 仕事	1.0	0.5
仕事は	0.5	1.0
事は仕	1.0	1.0
は仕事	1.0	0.5
仕事 #	0.5	1.0

FIG. 12C

APPEARANCE FREQUENCY OF 3-GRAM	
CHARACTER STRING	NUMBER OF TIMES
# 仕事	1
仕事は	1
事は仕	1
は仕事	1
仕事 #	1

FIG. 12E

1ST FACTOR 1.0
2ND FACTOR 0.0
PROBABILITY
CHARACTER STRING : # 仕 事 は 仕 事 #

⇧
POINTER

FIG. 12F

1ST FACTOR		1.0	0.5	1.0	1.0	0.0
2ND FACTOR	0.0	1.0	1.0	0.5	1.0	
PROBABILITY		1.0	0.5	0.5	1.0	← PRODUCT OF 1ST & 2ND FACTORS
CHARACTER STRING :	#	仕	事/	は/	仕	事 #
DIVISION POINT		▲		▲		

FIG. 13

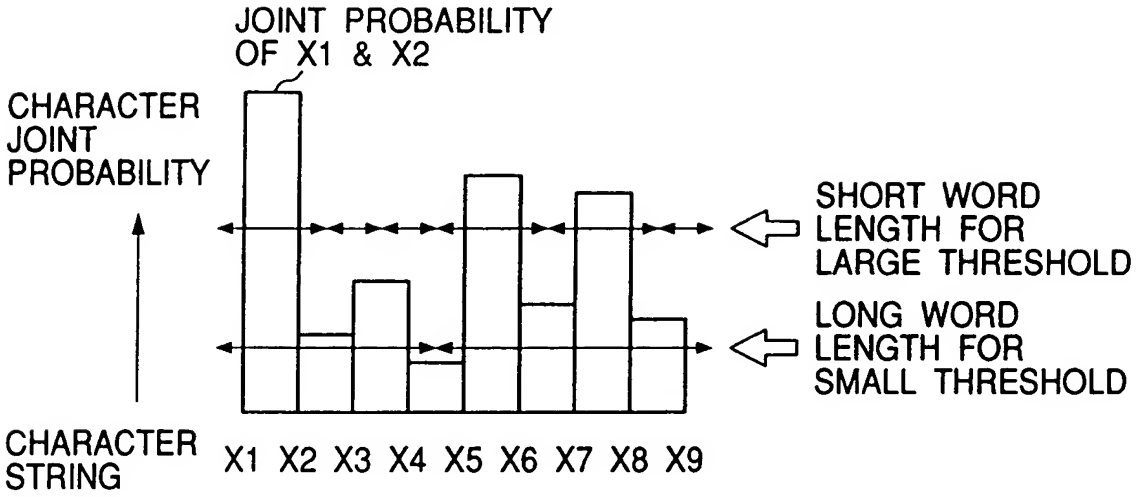

JOINT PROBABILITY
OF X1 & X2
CHARACTER
JOINT
PROBABILITY
SHORT WORD
LENGTH FOR
LARGE THRESHOLD
LONG WORD
LENGTH FOR
SMALL THRESHOLD
CHARACTER
STRING
X1 X2 X3 X4 X5 X6 X7 X8 X9

FIG. 14

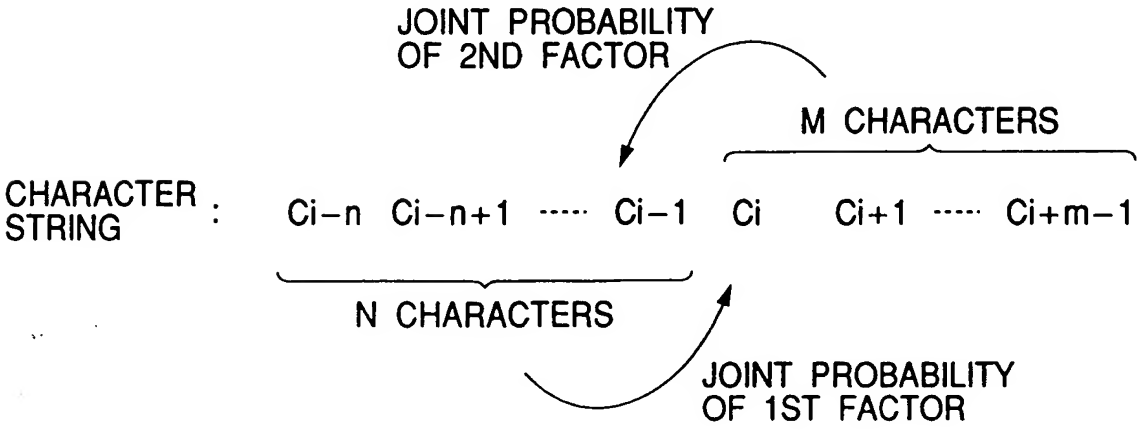

JOINT PROBABILITY
OF 2ND FACTOR
M CHARACTERS
CHARACTER
STRING :
Ci−n Ci−n+1 ····· Ci−1 Ci Ci+1 ····· Ci+m−1
N CHARACTERS
JOINT PROBABILITY
OF 1ST FACTOR

FIG. 15

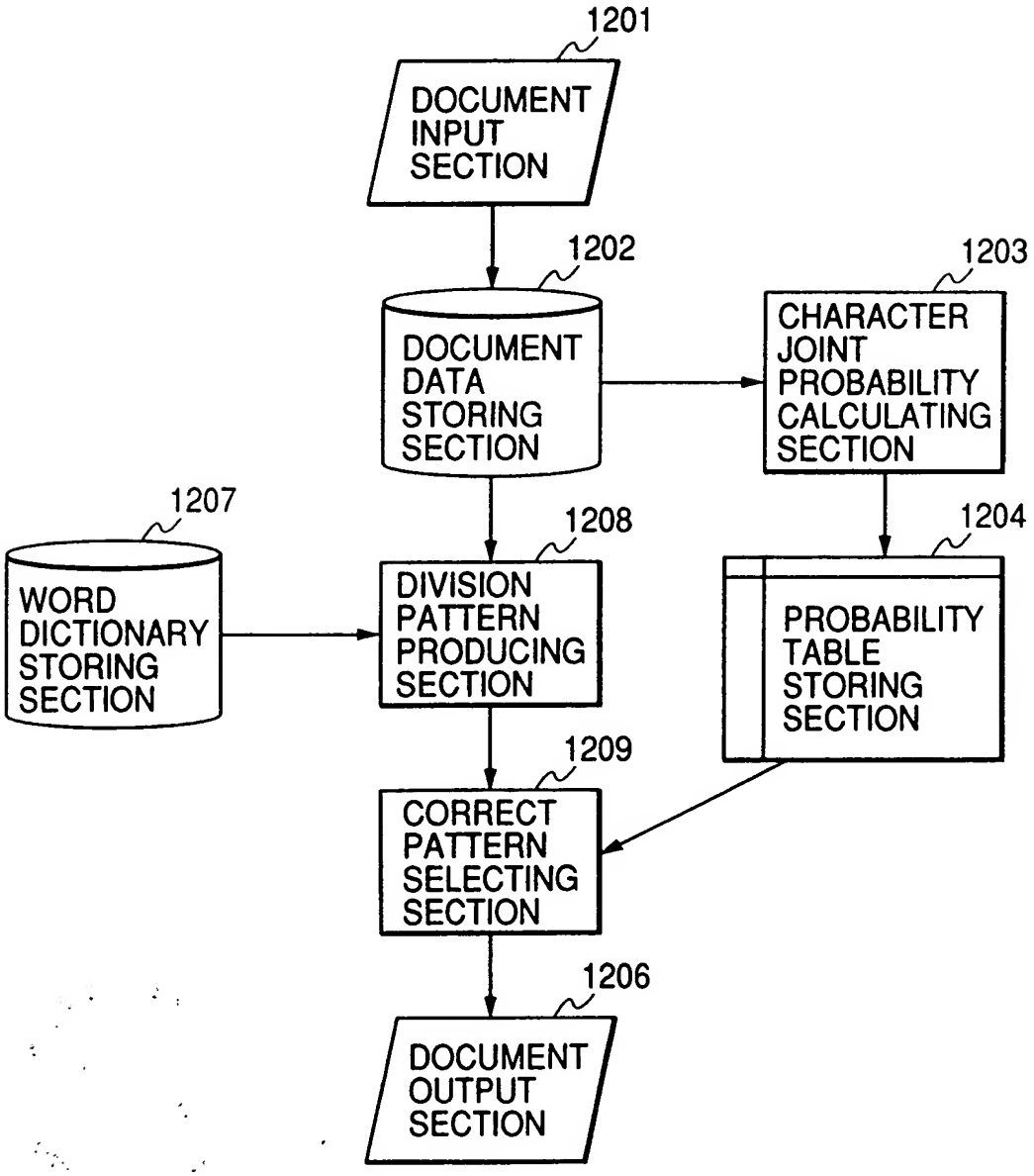

1201
DOCUMENT INPUT SECTION
1202
DOCUMENT DATA STORING SECTION
1203
CHARACTER JOINT PROBABILITY CALCULATING SECTION
1207
WORD DICTIONARY STORING SECTION
1208
DIVISION PATTERN PRODUCING SECTION
1204
PROBABILITY TABLE STORING SECTION
1209
CORRECT PATTERN SELECTING SECTION
1206
DOCUMENT OUTPUT SECTION

FIG. 16

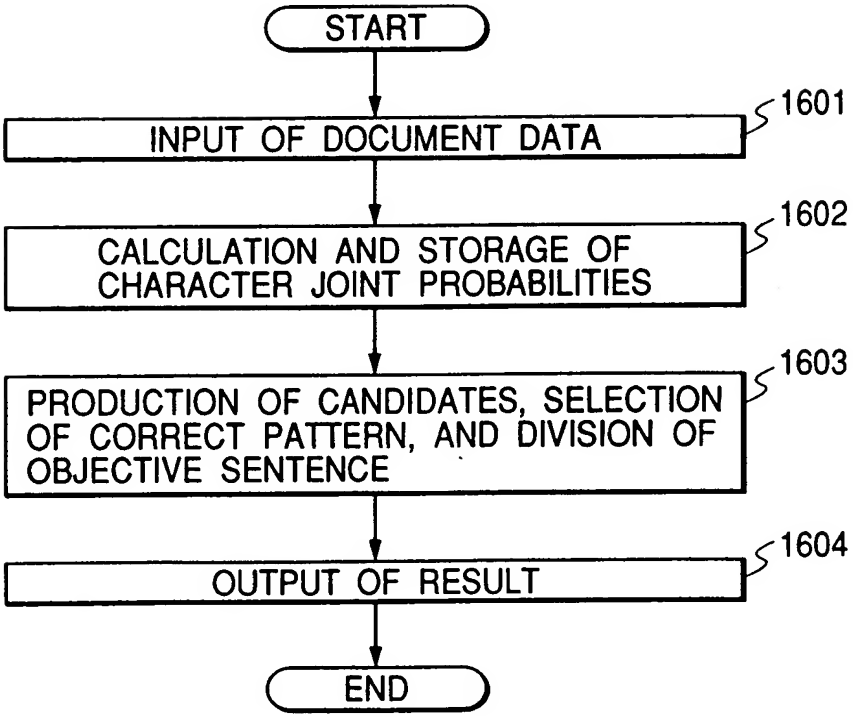

START
INPUT OF DOCUMENT DATA
1601
CALCULATION AND STORAGE OF CHARACTER JOINT PROBABILITIES
1602
PRODUCTION OF CANDIDATES, SELECTION OF CORRECT PATTERN, AND DIVISION OF OBJECTIVE SENTENCE
1603
OUTPUT OF RESULT
1604
END

FIG. 17A

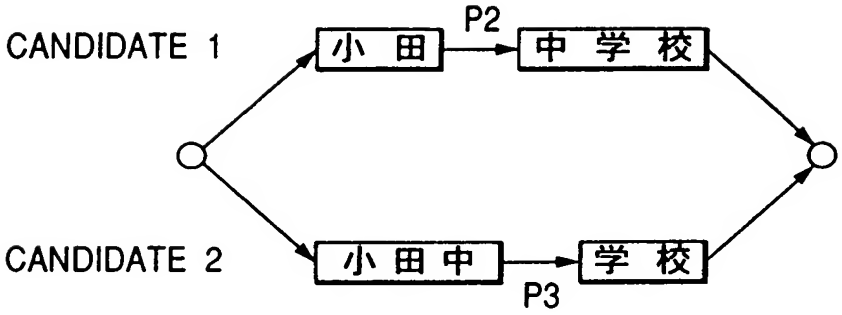


FIG. 17B

小田中学校 > P1=0.019
> P2=0.007
> P3=0.088
> P4=0.439

FIG. 18A

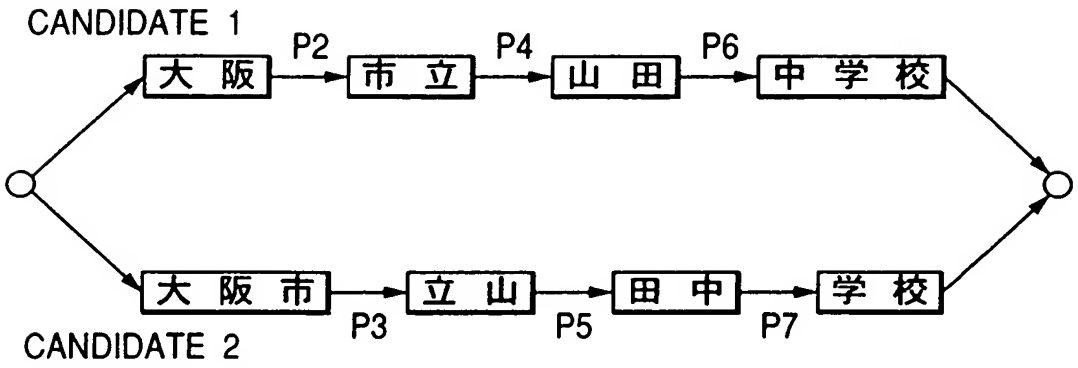


FIG. 18B

大阪市立山田中学校

> P1=0.599
> P2=0.141
> P3=0.174
> P4=0.006
> P5=0.111
> P6=0.006
> P7=0.088
> P8=0.439

FIG. 18C

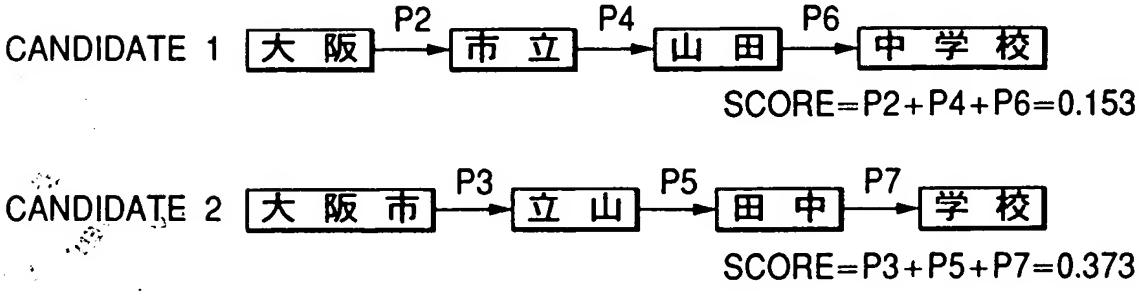


FIG. 19

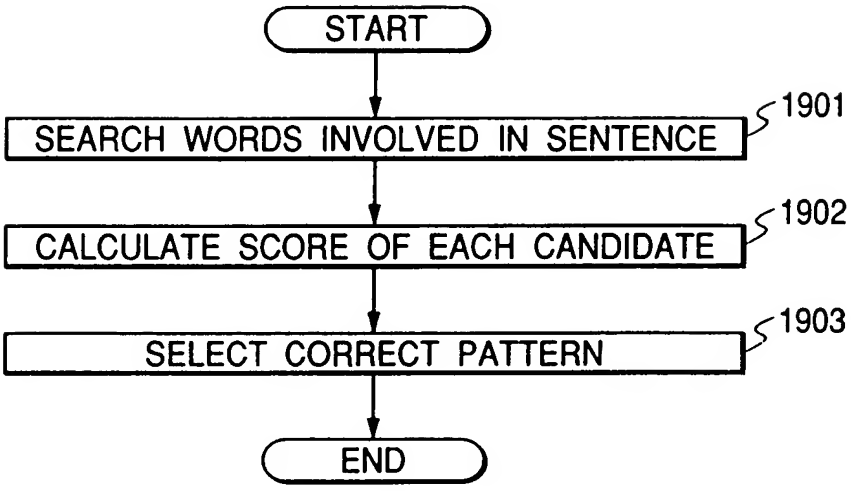

START
SEARCH WORDS INVOLVED IN SENTENCE
1901
CALCULATE SCORE OF EACH CANDIDATE
1902
SELECT CORRECT PATTERN
1903
END

FIG. 20

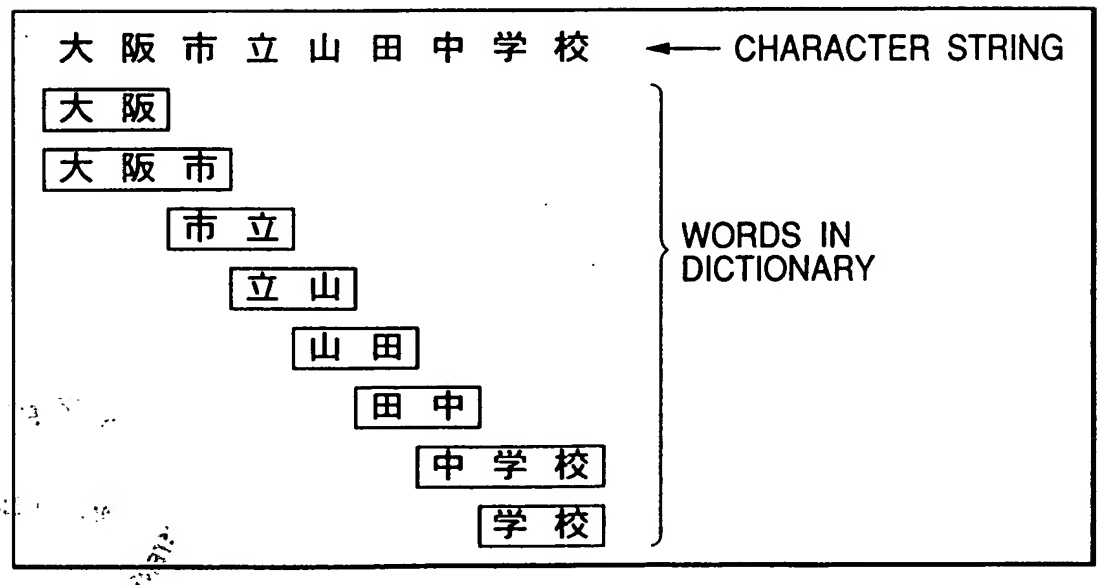

大 阪 市 立 山 田 中 学 校
CHARACTER STRING
大 阪
大 阪 市
市 立
立 山
山 田
田 中
中 学 校
学 校
WORDS IN DICTIONARY

FIG. 21

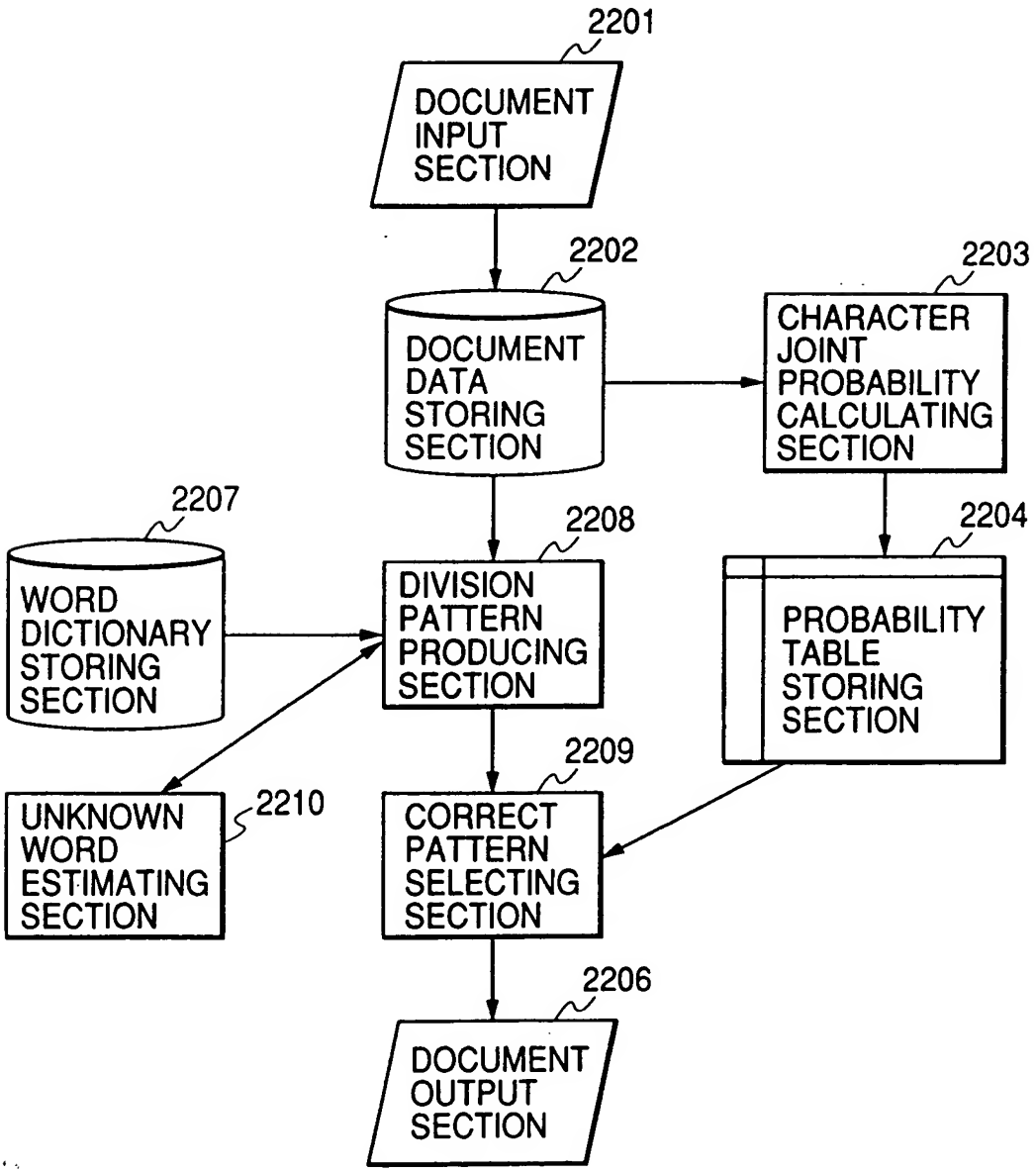

2201
DOCUMENT INPUT SECTION
2202
DOCUMENT DATA STORING SECTION
2203
CHARACTER JOINT PROBABILITY CALCULATING SECTION
2207
WORD DICTIONARY STORING SECTION
2208
DIVISION PATTERN PRODUCING SECTION
2204
PROBABILITY TABLE STORING SECTION
2209
UNKNOWN WORD ESTIMATING SECTION
2210
CORRECT PATTERN SELECTING SECTION
2206
DOCUMENT OUTPUT SECTION

FIG. 22

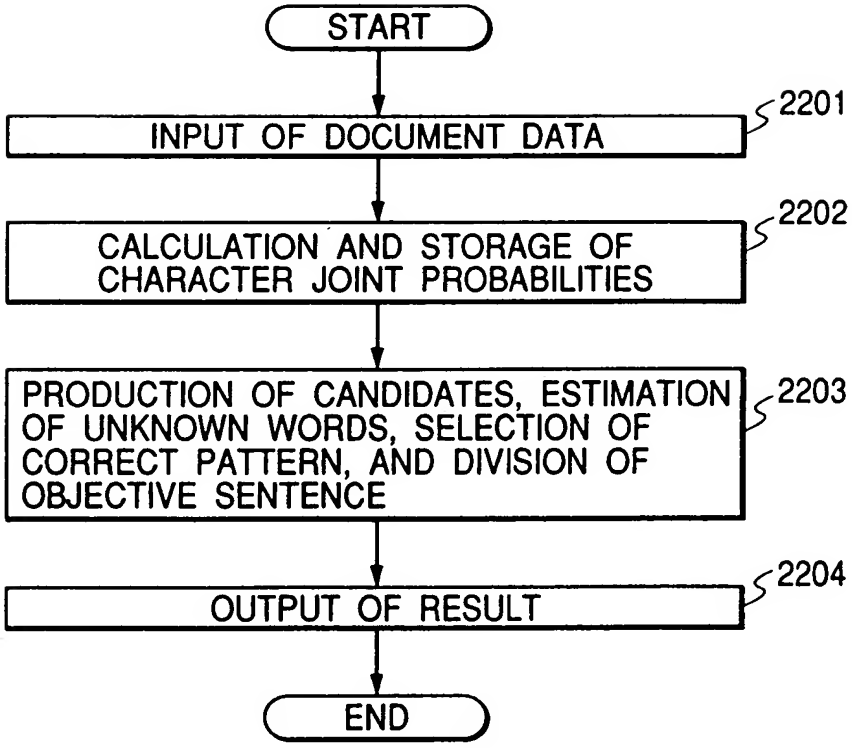

START
INPUT OF DOCUMENT DATA
2201
CALCULATION AND STORAGE OF CHARACTER JOINT PROBABILITIES
2202
PRODUCTION OF CANDIDATES, ESTIMATION OF UNKNOWN WORDS, SELECTION OF CORRECT PATTERN, AND DIVISION OF OBJECTIVE SENTENCE
2203
OUTPUT OF RESULT
2204
END

FIG. 23

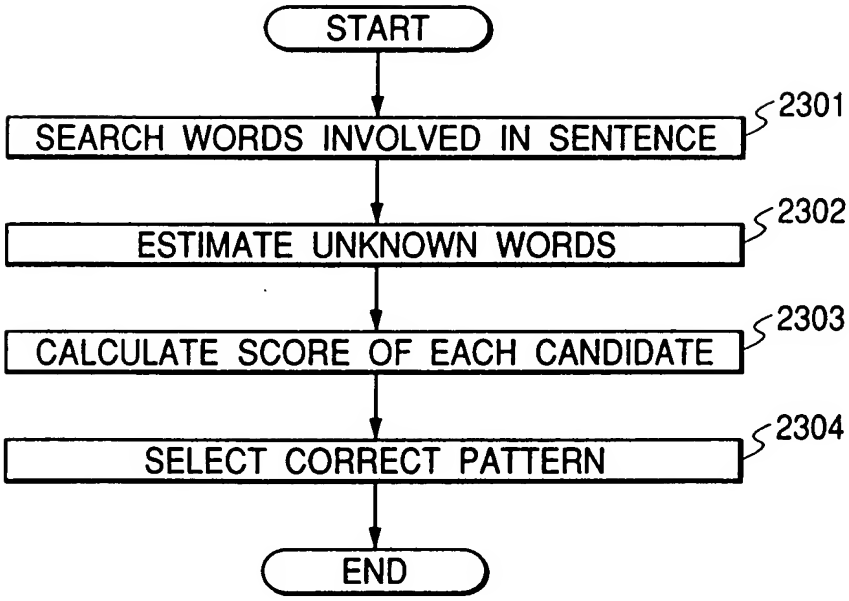

START
SEARCH WORDS INVOLVED IN SENTENCE
2301
ESTIMATE UNKNOWN WORDS
2302
CALCULATE SCORE OF EACH CANDIDATE
2303
SELECT CORRECT PATTERN
2304
END

FIG. 24

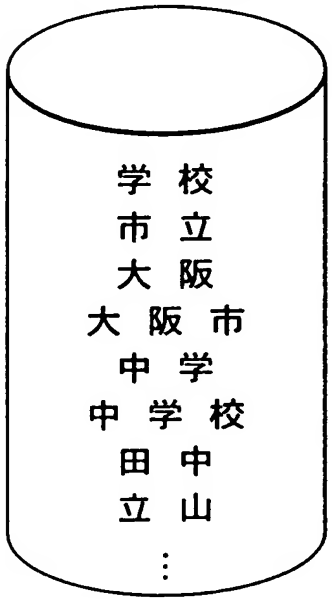

WORD DICTIONARY
学 校
市 立
大 阪
大 阪 市
中 学
中 学 校
田 中
立 山
⋮

FIG. 26

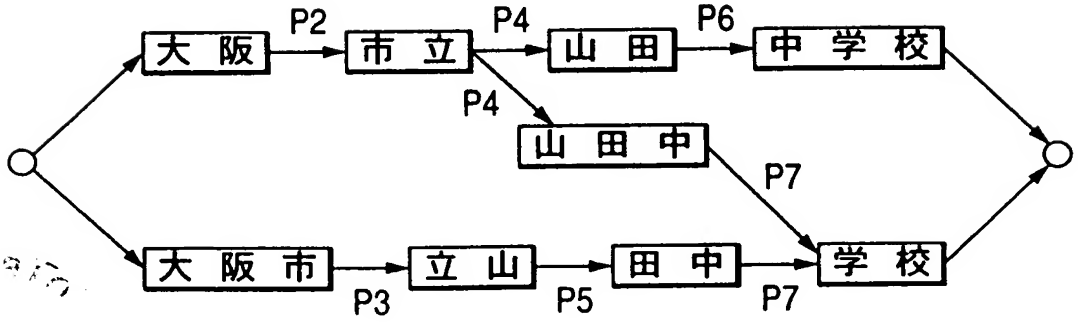

大 阪
P2
市 立
P4
山 田
P6
中 学 校
P4
山 田 中
P7
大 阪 市
P3
立 山
P5
田 中
P7
学 校

FIG. 27

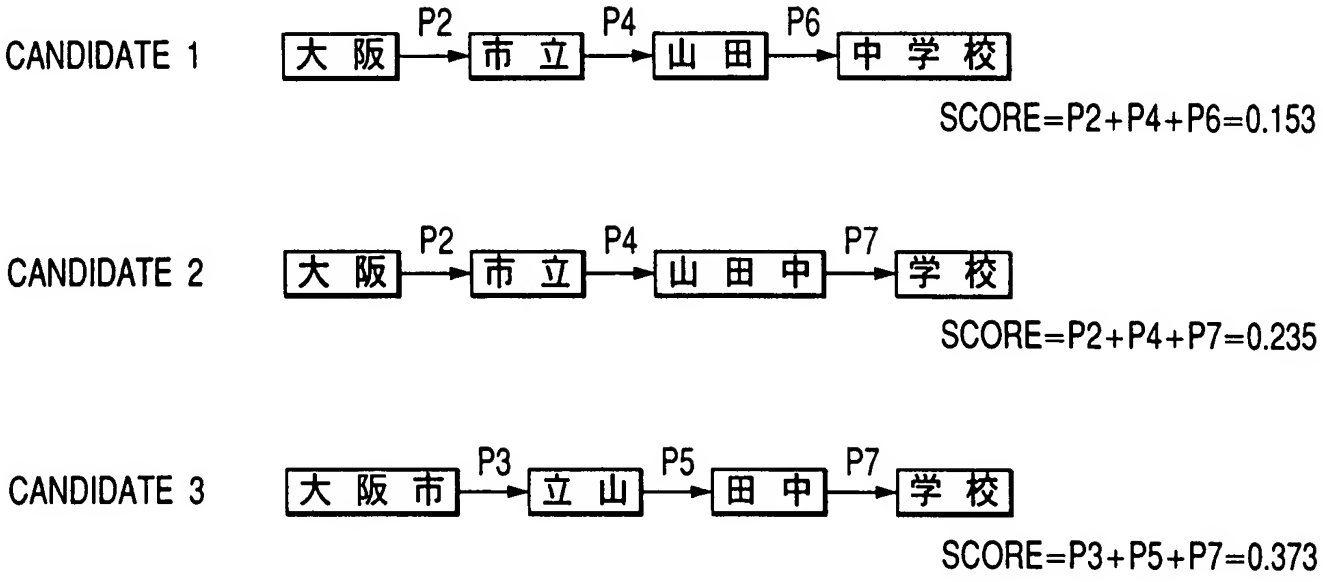

CANDIDATE 1
大 阪
P2
市 立
P4
山 田
P6
中 学 校
SCORE=P2+P4+P6=0.153
CANDIDATE 2
大 阪
P2
市 立
P4
山 田 中
P7
学 校
SCORE=P2+P4+P7=0.235
CANDIDATE 3
大 阪 市
P3
立 山
P5
田 中
P7
学 校
SCORE=P3+P5+P7=0.373

FIG. 28A

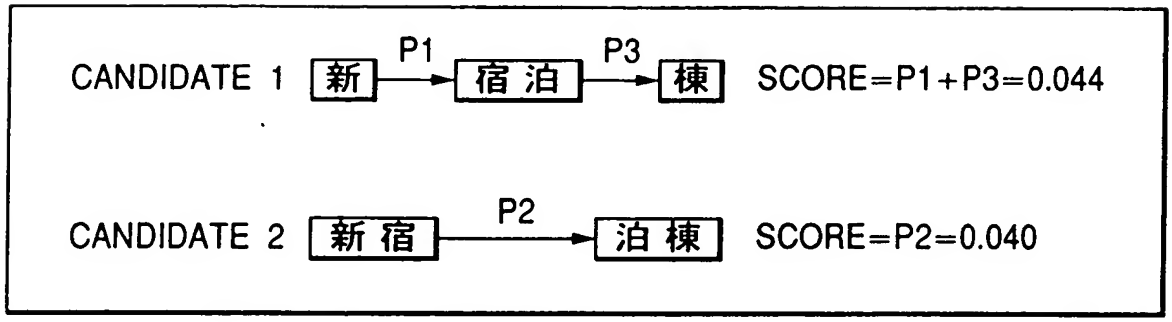


FIG. 28B

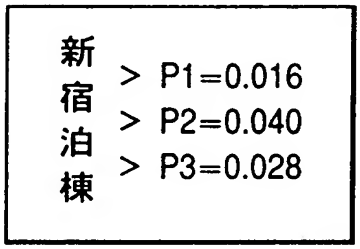
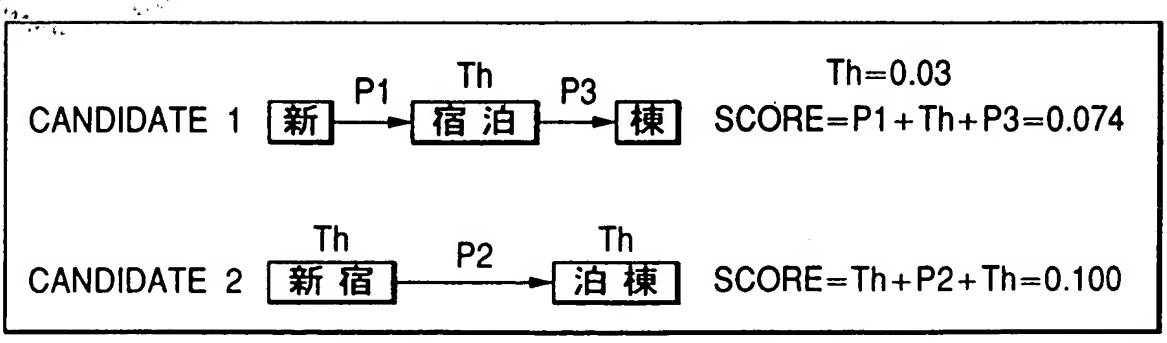


FIG. 28C



US 0974579504P1

Creation date: 08-07-2004
Indexing Officer: VWALKER - VIKITA WALKER
Team: OIPEBackFileIndexing
Dossier: 09745795

Legal Date: 07-12-2004

No.	Doccode	Number of pages
1	SRNT	6

Total number of pages: 6

Remarks:

Order of re-scan issued on .....